Received: 6 July 2019 | Revised: 6 February 2020 | Accepted: 10 February 2020
DOI: 10.1002/mpr.1822

ORIGINAL ARTICLE

WILEY

Objective measurement of attention deficit hyperactivity disorder symptoms outside the clinic using the QbCheck: Reliability and validity

Fredrik Ulberstad[1] | Hans Boström[1] | Mira-Lynn Chavanon[2] | Martin Knollmann[3] | James Wiley[4] | Hanna Christiansen[2] | Lisa B. Thorell[1]

[1]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[2]Clinical Child and Adolescent Psychology, Department of Psychology, Philipps University Marburg, Germany
[3]Department of Child and Adolescent Psychiatry, Psychosomatics, and Psychotherapy, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
[4]Focus-MD, Mobile, Alabama

Correspondence
Lisa B. Thorell, Department of Clinical Neuroscience, Karolinska Institutet, Nobels Väg 9, SE-171 77 Stockholm, Sweden.
Email: lisa.thorell@ki.se

Abstract
Objective measurements of ADHD symptom levels can be a highly valuable complement to ratings. However, sometimes it is not feasible to bring patients into the clinic/lab for assessment. The aim of the present study was therefore to evaluate the psychometric properties of the QbCheck, an online computerized test that measures errors and reaction time as well as activity during testing using the computer's built-in web camera. Study I (n = 27 adolescents/adults) investigated test–retest reliability and concurrent validity of the QbCheck. Study II included 142 adolescents/adults (69 with ADHD/73 controls) and investigated convergent and diagnostic validity, as well as usability, of the QbCheck. In Study I, the QbCheck showed high test–retest reliability and high concurrent validity. In Study II, high convergent validity was observed when studying associations between the QbCheck performed in the home and the QbTest performed at the clinic. In addition, the QbCheck discriminated well between patients with ADHD and controls, with a sensitivity of 82.6 and a specificity of 79.5. The QbCheck appears to be a valuable test with good psychometric properties and will thereby enable assessment of ADHD symptom levels in adolescents and adults outside the clinic in the home setting.

KEYWORDS
ADHD, diagnosis, objective measurement, reliability, validity

1 | INTRODUCTION

Attention deficit hyperactivity disorder (ADHD) is characterized by excessive levels of the following three behavioral symptoms: hyperactivity, impulsivity, and inattention. According to international guidelines (e.g., National Institute for Health and Care Excellence [NICE], 2018), the diagnostic procedures for ADHD should include an overall cognitive assessment, parent and teacher ratings (self-ratings if age allows), standardized clinical interviews, and if possible also observations in an educational or occupational setting. This approach is highly reliant on subjective measures, and previous studies have found only modest interrater agreement for ADHD symptom levels in adolescents and adults (e.g., Adler et al., 2008; Narad et al., 2015; Zucker, Morris, Ingram, Morris, & Bakeman, 2002). Using objective measurements of ADHD symptom levels could therefore be highly valuable in both research and clinical practice, and the Quantified Behavioural Test (QbTest; Ulberstad, 2016) has been introduced to meet this need. However, there is also a need to be able to collect objective measurements of ADHD symptom levels outside the clinic. The aim of the present study was therefore to examine the psychometric properties of the QbCheck, an online test that uses the built-in web camera found in most modern computers, thereby enabling easy testing of objective ADHD symptom levels outside the clinic.

Previous research has shown that assessment of ADHD symptom levels using ratings is likely to be influenced by various forms of bias related to, for example, rater characteristics such as age and gender (e.g., Schultz & Evans, 2012). Another type of bias is the halo effect, which is demonstrated, for example, by the fact that parents are more likely to rate their child as defiant when ADHD symptoms are present (e.g., DeVries, Hartung, & Golden, 2017). Rater bias is also evident when evaluating treatment effects, and the importance of using "blind" raters has therefore repeatedly been emphasized (e.g., Sonuga-Barke et al., 2013).

One way of dealing with the limitations of ratings is to include objective measurements of symptom levels in the assessment of ADHD. Different types of continuous performance tasks (CPT) have been used to assess inattention and impulsivity in both research and clinical settings (review by Hall et al., 2016). In the CPT test paradigm, the subject is presented with a rapid series of visual or auditory stimuli, and is instructed to respond as quickly as possible when a target stimulus is presented and to refrain from responding when a nontarget stimulus is presented. There are many different kinds of CPT paradigms. The present study used a task based on the "identical pairs paradigm," in which the participant is instructed to only respond when two consecutive stimuli are identical.

One limitation of the classical CPT paradigm is that it only measures two of the core symptoms of ADHD, inattention and impulsivity, but not hyperactivity. However, the QbTest (Ulberstad, 2016) includes a computerized CPT to measure inattention and impulsivity in combination with a high-resolution motion tracking system, which enables objective measurement of activity. Previous research has indicated that the QbTest has good test–retest reliability (Ulberstad, 2016) and that it can differentiate between patients with ADHD and healthy controls in childhood (Emser et al., 2018; Hult, Kadesjö, Kadesjö, Gillberg, & Billstedt, 2018), in adulthood (Edebol, Helldin, & Norlander, 2013; Emser et al., 2018; Hirsch & Christiansen, 2017; Lis et al., 2010), as well as in older adulthood (Bijlenga et al., 2019). QbTest scores have also been shown to be sensitive to medication effects (Bijlenga, Jasperse, Gehlhaar, & Sandra Kooij, 2015; Vogt & Williams, 2011). In addition, Vogt and Shameli (2011) examined the clinical utility of the QbTest by comparing children assessed using standard clinical assessment with or without the QbTest. The results showed that the ADHD diagnosis was less likely to be revised over a 12-month period in the group using the QbTest. Thus, objective measurement of ADHD symptom levels using the QbTest may reduce the risk of diagnostic errors.

In summary, previous research on the QbTest has shown that the test has good psychometric properties, that it can improve the diagnostic procedure, and that it can be useful in tracking treatment outcomes. However, performing the test at home can reduce the burden of in-clinic visits. Results from the *Assessing QbTest Utility in ADHD-Trial* (AQUA-Trial) showed that even though the QbTest reduced the appointment length by 15% and increased clinicians' confidence in their diagnostic decisions, cost savings where considered low (Hollis et al., 2018). Interestingly, a small qualitative study also conducted within the AQUA-Trial reported that clinicians thought that objective testing should ideally be performed *before* the initial appointment at the clinic (Hall et al., 2017). These findings indicate that both in terms of clinical utility and cost-effectiveness, there is a need for objective measurement of ADHD symptom levels outside the clinic. Besides being useful in the diagnostic process, objective measurement of ADHD symptom levels could also be valuable as part of screening programs or to more easily track changes in ADHD symptom levels during treatment. The aim of the present study was therefore to examine the psychometric properties of the QbCheck, a test that is substantially equivalent to the QbTest and that can be performed in the home or school setting. Instead of using a sophisticated motion-tracking device that needs to be installed at the clinic, the QbCheck is performed on a laptop and activity is recorded using the computer's built-in camera.

The present study included two separate studies. In Study I, we examined the test–retest reliability of the five variables derived from the QbCheck and the concurrent validity of the activity variable (i.e., Microevent) obtained in the QbCheck. In Study II, the primary objective was to examine diagnostic validity by investigating to what extent the QbCheck can be used to discriminate between adolescents/adults diagnosed with ADHD and healthy controls. In addition, convergent validity was examined by investigating the relation between the QbCheck performed in a domestic setting and the QbTest performed in the clinic. Finally, Study II also aimed to investigate the usability of the QbCheck. This includes measures of how many individuals had problems performing the test as well as the participants' experiences of performing the test on their own in the home setting.

2 | METHODS

2.1 | Participants and procedure

2.1.1 | Study I

In Study I, which aimed to evaluate both the test–retest reliability of the five QbCheck variables and the concurrent validity of the motion detection variable, both healthy controls and individuals diagnosed with ADHD were included. In total, 27 individuals participated in Study I. However, two participants were excluded from the analysis, one due to technical errors and one due to disengagement during the test. The sample included in the analyses consisted of 11 patients diagnosed with ADHD (54.5% males) and 14 healthy controls (57.1% males). The participants ranged in age from 12 to 59 years, with a mean age of 29.42 years (*SD* = 16.47) for the patients with ADHD and 23.93 years (*SD* = 13.80) for the controls.

The healthy controls were recruited via convenience sampling, schools and advertisement in Marburg, Germany. Participants diagnosed with ADHD were recruited via the ADHD outpatient clinic of Philipps University, Marburg, Germany. All participants in the clinical sample were shown to meet the diagnostic criteria according to DSM-5 (American Psychiatric Association [APA], 2013) assessed using a standardized procedure at the clinic. After giving their written informed consent to participate in the study, the participants performed the test twice in a clinical environment, with the second test being performed 1–2 hr after the first. At both test and retest assessment, a QbTest camera was installed to enable simultaneous measurement of activity using both the computer's built-in camera and the infrared QbTest camera (i.e., concurrent validity).

2.1.2 | Study II

In Study II, which aimed to evaluate the validity of the different test variables, a total of 142 participants between 12 and 60 years of age were included. More specifically, the sample consisted of 69 individuals with ADHD (52.2% males; mean age = 27.58 [*SD* = 12.12]) and 73 healthy controls (43.8% males; mean age = 26.16 [*SD* = 9.55]). When comparing the group with ADHD and the controls, no significant differences were found with regard to age, $t = .78$, $p = .44$, or sex, $\chi^2 = .99$, $p = .32$. The inclusion criteria for the group with ADHD and the controls were the same as in Study I. Ratings of ADHD symptom levels were completed by parents for the adolescents and by the participant himself/herself for adults. One adult was excluded from the control group due to high levels (i.e., 5 symptoms) of either inattention or hyperactivity/impulsivity according to the symptom criteria listed in DSM-5 (APA, 2013).

The data from the healthy controls were collected in collaboration with Philipps University, Marburg, Germany and the Karolinska Institutet, Stockholm, Sweden via convenience sampling, at schools and workplaces. Individuals diagnosed with ADHD were recruited from outpatient clinics associated with Philipps University Marburg, Germany; University of Duisburg-Essen, Germany; Karolinska Institutet, Stockholm Sweden; and the Focus-MD group, Mobile, Alabama. After providing their written informed consent, participants were asked to perform the QbCheck at home before arriving at the clinic for their assessment, which included the QbTest.

2.2 | Materials

2.2.1 | QbTest and QbCheck

As described in the introduction, the QbTest (Ulberstad, 2016) is a computerized test aimed at measuring inattention and impulsivity in combination with a high-resolution motion tracking system designed to provide an objective measurement of activity. The QbCheck is an online version of the QbTest created for easy use in the home or school setting. Rather than using an advanced motion tracking system (i.e., an infrared camera), the QbCheck uses the built-in web camera in any ordinary computer to track activity. Instead of the response button used in the QbTest, the QbCheck uses the computer's space bar to record responses. Both the QbTest and the QbCheck include two different versions with different cognitive load, which are used for different age groups. However, only the adolescent/adult version was included in the present study. In this version, the participants are instructed to respond as quickly as possible when a target is presented on the computer screen, but to refrain from responding when a nontarget is presented. Four types of stimuli are included (i.e., a red circle, a blue circle, a red square, and a blue square), and the test is based on the unconditional identical pairs principle, which means that a stimulus is defined as a target if it is identical in shape and color to the stimulus immediately preceding it. The stimuli are presented for 200 ms with a stimuli-onset interval of 2 s. The total number of stimuli is 600, with 25% being targets. The order of the targets and nontargets is randomized to prevent learning effects over repeated test administrations.

Five principle variables, reflecting different markers associated with the three cardinal symptoms of ADHD (i.e., hyperactivity, inattention, and impulsivity), are derived from the QbCheck. A more detailed description of the different variables included in the QbCheck is presented in Table 1. In the present study, we used Q-scores with a mean of 0 and a *SD* of 1. These scores can be interpreted as similar to Z-scores and are derived by comparing the participant's raw scores to the normative data matched for age and sex using the normative QbTest database. More information on the QbTest and the normative groups can be found in the QbTest technical manual (Ulberstad, 2016).

Because of the QbCheck is administered in the home setting, more comprehensive instructions are provided within the test compared to the QbTest. This includes a step-by-step guide using computer animations in combination with voice/text instruction describing both environmental requirements (e.g., turning phones off, checking the lighting in the room, what type of chair to sit on)

TABLE 1 Description of the variables included in the QbCheck

QbCheck variables	
Activity	
Microevents	Number of position changes of the head (≥ 1 mm) on the x-axis
Attention	
Omission errors	Not responding to a target (% of total number of targets)
Reaction time	Time to respond for correct responses
Reaction time variability	*SD* of the reaction time
Impulsivity	
Commission errors	Responding to a nontarget (% of total number of nontargets)

and specific test instructions. The test cannot be started unless the participant has completed all steps of the instructions. A practice test is also included during which the motion tracking is evaluated and the test can only be started if the motion tracking is of high quality.

2.2.2 | Usability of the QbCheck

User friendliness was assessed using a brief questionnaire completed by the participant or the participant's parent in case he/she was below age 15. Ratings were made on a scale from 1 to 10, with higher values indicating higher user-friendliness. The following three questions were included: First, how easy was the QbCheck to use? Second, how easy was it to understand and perform the preparations before starting the test? Third, how easy was it to understand and follow the test rules during the test? In addition, the participants were asked (yes/no question) whether they experienced any technical problems with the computer/camera check. To obtain more detailed data, participants were also able to provide more detailed descriptions of possible problems in relation to each one of the questions and to give suggestions for improvement.

2.3 | Statistical analyses

All analyses were performed using IBM SPSS Statistics version 25 (SPSS Co., Chicago). The presence of outliers was first examined using the Outlier Labeling Criteria (Hoaglin & Iglewicz, 1987), but no outliers were found. In Study I, test–retest reliability was first analyzed using intra-class correlations (ICC; two-way mixed model, absolute agreement). Second, correlations were used to investigate the association between the activity variable (i.e., Microevents) obtained from the built-in camera included in the QbCheck and simultaneous measurement of activity using the infrared QbTest camera. Data from both the test and retest assessment were included in these analyses.

In Study II, different aspects of validity of the QbCheck were investigated. Convergent validity was examined by investigating correlations between the QbCheck completed in the home setting and the corresponding variables obtained from the QbTest completed at the clinic. With regard to diagnostic validity, multivariate analysis of variance (MANOVA) and follow-up *t* tests were first of all used to compare individuals with ADHD and controls with regard to the five QbCheck variables. Effect sizes were calculated using Hedges's *g*, where 0.20 is considered a small effect, 0.50 a medium-sized effect and 0.80 a large effect (Cohen, 1988). Next, logistic regression was used to examine how well the QbCheck variables could classify individuals and controls. Results are reported in terms of measures' sensitivity (i.e., correctly identified individuals with ADHD) and specificity (i.e., correctly identified controls) as well as Receiver Operating Characteristics (ROC) curves. Finally, with regard to usability, we used either *t* tests (the dimensional variables) or chi-square test (the categorical variable) to compare the group with ADHD to the controls.

3 | RESULTS

3.1 | Study I

In the first part of the study when assessing test–retest reliability, the results showed excellent or near excellent intra-class correlations for all variables: 0.90 for Microevents, 0.84 for Omission Errors, 0.82 for Commission Errors, 0.96 for Reaction Time, and 0.88 for Reaction Time Variability. With regard to concurrent validity, results showed that the activity variable Microevents registered through the built-in camera included in the QbCheck was very highly correlated with Microevents simultaneously assessed using the infrared camera included in the QbTest, $r = .91$ $p < .001$.

3.2 | Study II

3.2.1 | Convergent validity

In the part of the study assessing convergent validity, data were only available for the ADHD group. In addition, two participants in the group with ADHD did not complete the QbTest and therefore had to be excluded from the analyses, resulting in a total sample size of $N = 67$. The results (see Table 2) showed that correlations between the QbCheck and the corresponding variables obtained for the QbTest were high for all five variables (*r*s ranging between 0.50 and 0.68).

3.2.2 | Diagnostic validity

First, group differences between individuals with ADHD and healthy controls were investigated. An overall MANOVA showed a significant main effect of group, F (1, 140) = 22.83, $p < .001$. Descriptive data and results of post hoc *t* tests are presented in Table 3. The results showed that individuals with ADHD performed more poorly

TABLE 2 Estimates of reliability and validity for the five QbCheck variables

	Reliability: test–retest (n = 25)	Diagnostic validity: ROC curves (n = 142)	Convergent validity: correlations with the QbTest (n = 67)
Microevents	0.90	0.80	0.52***
Omission errors	0.84	0.75	0.74***
Commission errors	0.82	0.74	0.50***
Reaction time	0.96	0.73	0.68***
Reaction time variability	0.88	0.81	0.65***

***$p < .001$.

TABLE 3 Q-Score means (M), *SD*, results of *t* test and effect sizes (ES) when comparing individuals with ADHD and controls on the five QbCheck variables

	ADHD group M (*SD*) *n* = 69	Control group M (*SD*) *n* = 73	*t*-value	ES
Hyperactivity				
Microevents	2.24 (1.31)	0.71 (.92)	7.97***	1.36
Inattention				
Omission errors	1.28 (1.15)	0.31 (.98)	5.42***	0.90
Reaction time	0.95 (1.13)	0.13 (.83)	4.89***	0.83
Reaction time variability	1.55 (1.06)	0.28 (.91)	7.64***	1.29
Impulsivity				
Commission errors	1.59 (1.12)	0.63 (1.17)	5.11***	0.84

***$p < .001$.

compared to controls on all five QbCheck variables. Effect sizes for the five QbCheck variables were all above the limit of what is normally considered a large effect (i.e., ≥ 0.80).

The logistic regression analyses were also significant, $\chi^2 = 85.53$, $p < .001$, $R^2 = 0.60$. The five QbCheck variables could correctly classify 115 (81.0%) of the participants with a sensitivity of 82.6 and a specificity of 79.5. ROC curves for the different variables are shown in Figure 1. The variables with the highest AUC were Microevent (0.82) and Reaction Time Variation (0.81). The values for the AUC for Omission Errors (0.76), Commission Errors (0.75) and Reaction Time (0.73) were somewhat lower, but still within the range of what is normally considered fair (see Table 2).

3.2.3 | Usability of the QbCheck

With regard to potential technical problems performing the QbCheck, results showed that 149 participants managed to initiate a test. Of these, 142 participants completed the test, two participants experienced technical problems with the camera, and four participants ended the test in the middle of the session for unknown reasons. In addition, one participant intentionally did not follow through the test. Of the participants who did not manage to complete the test, all but one belonged to the group with ADHD.

Results for the three questions assessing the usability of the QbCheck (see Table 4) showed that mean values were high (all mean values ≥8.06 on a 10-point scale). The participants in the control group found it easier to understand and perform the preparations before starting the test compared to the participants in the group with ADHD, although it should be noted that mean values were high for both groups. There were no significant group differences for the other two feasibility questions. By far the most common reason for a score lower than eight on ease of using the QbCheck was that the test was experienced as taking a long time and that it was therefore difficult to stay focused during the test.

Finally, the participants also answered a yes/no question asking whether they had any problems with the computer/camera check. In total, 30 (24%) participants (30.4% in the group with ADHD and 16.1% in the control group) experienced some sort of technical problem. There was a tendency for the group with ADHD to experience more technical problems compared to the control group, $\chi^2(1, N = 125) = 3.50$, $p = .06$. When asked to describe these problems in more detail, very few reported any serious technical problems and most appeared to be related to not having read the instructions carefully enough. The most commonly reported problems were using an Internet browser that was not compatible with the QbCheck (e.g., Mac users have to use a browser other than Safari), having a generally poor Internet connection (i.e., they reported often having a poor connection, not just when using the QbCheck), or having problems with calibrating the computer's built-in camera for use with the QbCheck.

4 | DISCUSSION

The aim of the present study was to examine the psychometric properties of the QbCheck, a new test for obtaining objective measurement of ADHD symptom levels in the home setting. The results showed that the QbCheck has good test–retest reliability. In addition, the QbCheck has good concurrent and convergent validity when studying correlations to corresponding variables obtained from the QbTest; it also has good diagnostic validity for discriminating between individuals with ADHD and healthy controls. Finally, of the individuals who managed to initiate the test, only a small minority (4.7%) was not able to produce usable data, and most participants thought instructions were easy to understand and follow.

By showing that the QbCheck has good psychometric properties, with high test–retest reliability and high validity, the present study introduces a new assessment method that could be seen as an important complement to rating scales and clinical interviews, with the added benefit of being a cost-effective method of assessing ADHD symptom levels outside the clinic. Assessing symptom levels using some sort of objective measurement tool should be considered important because ratings are always greatly influenced by personal biases (cf. Emser et al., 2018). However, it has been argued that tests and ratings capture at least partly different constructs and that tests

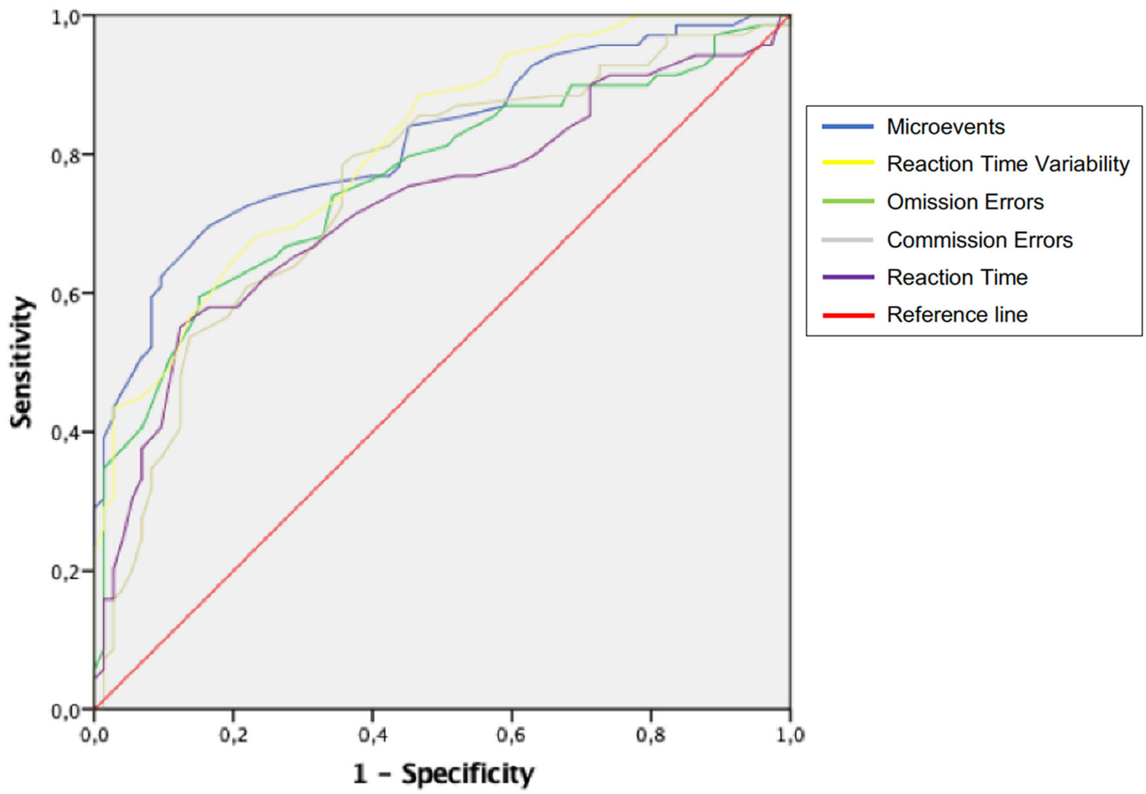

FIGURE 1 Receiver operating characteristics curves (ROC) for the QbCheck variables

TABLE 4 Means, standard deviations, results of independent *t* test, and effect sizes (d) to assess differences between means of the ADHD group and the control group on questions of usability

	ADHD group M (*SD*) *n* = 56[a]	Control group M (*SD*) *n* = 69[a]	*t*-value	ES
How easy was the QbCheck to use?	8.21 (2.26)	8.06 (2.04)	0.41	0.07
How easy was it to understand and perform the preparations before starting the test?	8.89 (1.89)	9.61 (1.05)	2.53*	0.47
How easy was it to understand and follow the test rules during the test?	9.13 (1.64)	9.39 (1.39)	1.03	0.17

Note: All questions were assessed on a scale from 0 to 10.
*$p < .05$.
[a]Thirteen missing for the ADHD group and four missing for the controls.

should therefore be seen as a complement to rather than as a replacement for ratings (Toplak, West, & Stanovich, 2013). It could be argued that use of objective measures of ADHD symptom levels is especially important in adult ADHD, as ratings made by a significant other (i.e., a partner or parent) are not always available and clinicians might therefore have to rely solely on self-ratings.

With regard to concurrent and convergent validity, the present study demonstrated high correlations between the QbCheck and the QbTest. Interestingly, high construct validity was demonstrated for all three symptom domains (i.e., inattention, hyperactivity, and impulsivity). However, this does not necessarily mean that home testing should replace testing at the clinic. Rather, these findings should be taken as an indication that there might be instances where the QbCheck can be used to obtain data on objective ADHD symptom levels when it is not feasible to bring participants/patients to the clinic (e.g., when collecting repeated measurements for tracking treatment progress). Although this needs to be examined in future research, the fact that the QbCheck can reduce the burden of in-clinic visits could make it a valuable screening instrument or complement in the diagnostic process. As stated in the introduction, a previous study (Hollis et al., 2018) showed that clinicians who used the QbTest found it to be a good complement to other methods in the process of diagnosing ADHD. Future research should also examine if there are situations in which a clinic-administered QbTest is preferable to the home-administered QbCheck. Possibly, the QbCheck should not be used if the test-taker has very limited experience using a laptop and does not have access to help within the family. In addition, a clinic-administered QbTest is preferable if there are reasons to believe that the home setting is too chaotic or if one might suspect that someone else than the intended test-taker will perform the test.

With regard to diagnostic validity, the QbCheck was able to discriminate relatively well between adolescents/adults with ADHD and controls, with a sensitivity of 82.6% and a specificity of 79.5%. These results are similar to those obtained in previous studies in adults and children examining the ability to discriminate between adults with ADHD and healthy controls using the QbTest in a clinical setting (Edebol et al., 2013; Emser et al., 2018). Interestingly, Microevents and Reaction Time Variability were the two variables with the highest AUC. Several previous studies using the QbTest have also found that measures of hyperactivity and inattention, rather than impulsivity, were best at differentiating between ADHD and other clinical diagnoses (Hult et al., 2018; Söderström, Pettersson, & Nilsson, 2013). Altogether, these findings clearly show that tracking the movements of patients with ADHD adds valuable information in addition to that captured through use of most CPT paradigms, which only register errors and reaction time.

The other variable that showed a high AUC was Reaction Time Variability. This variable is usually regarded as a measure of inattention, capturing how well the participant can regulate his/her internal state during a task. The fact that individuals with ADHD have higher Reaction Time Variability than controls is one of the most replicated findings in this area of research (review by Castellanos et al., 2005).

Finally, usability analyses showed that only a very small proportion of participants who logged into the system were not able to complete the test and generate usable data. Few technical problems were experienced, most of which could be solved with only slight changes in the test instructions (e.g., clarifying what Internet browsers are compatible with the QbCheck). It is also important to acknowledge that the most common reason for reporting a low score regarding ease of using the QbCheck was that it was difficult to stay focused during the entire task. However, as the purpose of the whole task is to assess attention, it should be considered a good thing that some individuals found the task challenging. It will be important for future research to investigate whether a shortened version of the QbCheck could be a valuable tool even for adolescents and adults, especially as one important purpose of the QbCheck could be to obtain repeated assessments of ADHD symptom levels as part of, for example, tracking treatment progress.

4.1 | Limitations, future directions, and conclusions

One important limitation of the present study is that it only included the adolescent/adult version of the QbCheck. As with the QbTest, a simplified version of the test for children aged 6–12 years is available, and it will be important for future research to examine the psychometric properties of this version of the QbCheck. Establishing a QbCheck version for children could enable objective measurement of ADHD symptoms as part of early screening for ADHD, for use either at home or in the school setting. In addition, the usability data in the present study only included participants who managed to initiate a test. Thus, it is not known how many participants logged into the system but failed to start the test. Another limitation worth mentioning is that test–retest reliability was examined within a clinical setting instead of the home setting.

As for future studies, it will be important to examine to what extent the QbCheck can be a valuable tool for diagnosing ADHD and monitoring symptom levels over time. It has been suggested that objective measurement of ADHD symptom levels may be important both during initial titration and for tracking patients' progress during the treatment (Ogundele, Ayyash, & Banerjee, 2011). In line with this suggestion, previous research has shown that the QbTest is more sensitive to medication effects than ratings are (Bijlenga et al., 2015). In the future, it could be possible to visualize the results of the QbCheck and make this information available to clinicians and/or the patients themselves, thereby not relying solely on subjective symptom reports for tracking treatment effects. More frequent measurement of the patient's ADHD symptom levels could also increase treatment adherence if the clinician allows the patient himself/herself to access the report generated from the QbCheck. However, this will probably call for a weighted summary score that can be interpreted more easily. In conclusion, the present results indicate that the adolescent/adult version of the QbCheck has good psychometric properties and that the QbCheck may therefore be an important contribution to the area of digital health, as it would allow clinicians to obtain objective measures of ADHD symptom levels without bringing patients into the clinic.

ACKNOWLEDGMENTS

The authors wish to thank Jessica Lindgren for her valuable help with the data collection.

CONFLICT OF INTEREST

Fredrik Ulberstad and Hans Boström are employed at Qbtech AB, the company that markets both the QbTest and the QbCheck.

ORCID

Lisa B. Thorell https://orcid.org/0000-0002-7417-6637

REFERENCES

Adler, L. A., Faraone, S. V., Spencer, T. J., Michelson, D., Reimherr, F. W., Glatt, S. J., … Biederman, J. (2008). The reliability and validity of self- and investigator ratings of ADHD in adults. *Journal of Attention Disorders*, *6*, 711–719.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Bijlenga, D., Jasperse, M., Gehlhaar, S. K., & Sandra Kooij, J. J. S. (2015). Objective QbTest and subjective evaluation of stimulant treatment in adult attention deficit-hyperactivity disorder. *European Psychiatry: The Journal of the Association of European Psychiatrists*, *30*(1), 179–185. https://doi.org/10.1016/j.eurpsy.2014.06.003

Bijlenga, D., Ulberstad, F., Thorell, L. B., Christiansen, H., Hirsch, O., & Kooij, J. J. S. (2019). Objective assessment of ADHD in older adults compared to controls using the QbTest. *International Journal of Geriatric Psychiatry*, *34*, 1526–1533.

Castellanos, F. X., Sonuga-Barke, E. J., Scheres, A., Di Martino, A., Hyde, C., & Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intraindividual variability. *Biological Psychiatry*, *57*, 1416–1423.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

DeVries, L. N., Hartung, C. M., & Golden, T. L. (2017). Negative halo effects in parent ratings of ADHD and ODD. *Journal of Psychopathology and Behavioral Assessment*, *39*(2), 179–188.

Edebol, H., Helldin, L., & Norlander, T. (2013). Measuring adult attention deficit hyperactivity disorder using the quantified behavior test plus. *Psychological Journal*, *2*, 48–62.

Emser, T. S., Johnston, B. A., Steele, J. D., Kooij, J. J. S., Thorell, L. B., & Christiansen, H. (2018). Assessing ADHD symptoms in children and adults: Evaluating the role of objective measures. *Behavioral and Brain Functions*, *14*, 11. https://doi.org/10.1186/s12993-018-0143-x

Hall, C. L., Valentine, A. Z., Groom, M. J., Walker, G. M., Sayal, K., Daley, D., & Hollis, C. (2016). The clinical utility of the continuous performance test and objective measures of activity for diagnosing and monitoring ADHD in children: A systematic review. *European Child & Adolescent Psychiatry*, *25*, 677–699. https://doi.org/10.1007/s00787-015-0798-x

Hall, C. L., Valentine, A. Z., Walker, G. M., Ball, H. M., Cogger, H., Daley, D., … Hollis, C. (2017). Study of user experience of an objective test (QbTest) to aid ADHD assessment and medication management: A multi-methods approach. *BMC Psychiatry*, *17*(1), 66.

Hirsch, O., & Christiansen, H. (2017). Factorial structure and validity of the quantified behavior test plus (Qb+@). *Assessment*, *24*, 1037–1049.

Hoaglin, D. C., & Iglewicz, B. (1987). Fine tuning some resistant rules for outliers labeling. *Journal of the American Statistical Association*, *82*, 1147–1149.

Hollis, C., Hall, C. L., Huo, B., James, M., Boadu, J., … Morris, R. (2018). The impact of a computerised test of attention and activity (QbTest) on diagnostic decision-making in children and young people with suspected attention deficit hyperactivity disorder: Single-blind randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *59*, 1298–1308.

Hult, N., Kadesjö, J., Kadesjö, B., Gillberg, C., & Billstedt, E. (2018). ADHD and the QbTest: Diagnostic validity of QbTest. *Journal of Attention Disorders*, *22*, 1074–1080.

Lis, S., Baer, N., Stein-en-Nosse, C., Gallhofer, B., Sammer, G., & Kirsch, P. (2010). Objective measurement of motor activity during cognitive performance in adults with attention-deficit/hyperactivity disorder. *Acta Psychiatrica Scandinavica*, *122*, 285–294.

Narad, M. E., Garner, A. A., Peugh, J. L., Tamm, L., Antonini, T. N., Kingery, K. M., … Epstein, J. N. (2015). Parent–teacher agreement on ADHD symptoms across development. *Psychological Assessment*, *27*, 239–248.

National Institute for Health and Care Excellence. (2018). Attention deficit hyperactivity disorder (NICE Quality Standard No. 39). Retrieved from https://www.nice.org.uk/guidance/qs39

Ogundele, M. O., Ayyash, H. F., & Banerjee, S. (2011). Role of computerised continuous performance task tests in ADHD. *Progress in Neurology and Psychiatry*, *15*(3), 8–13. https://doi.org/10.1002/pnp.198

Schultz, B. K., & Evans, S. W. (2012). Sources of Bias in teacher ratings of adolescents with ADHD. *Journal of Educational and Developmental Psychology*, *2*, 151–162.

Söderström, S., Pettersson, R., & Nilsson, K. (2013). Quantitative and subjective behavioural aspects in the assessment of attention-deficit hyperactivity disorder (ADHD) in adults. *Nordic Journal of Psychiatry*, *68*, 30–37.

Sonuga-Barke, E. J. S., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., … Sergeant, J. (2013). Nonpharmacological interventions for ADHD: Systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *American Journal of Psychiatry*, *170*, 275–289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*, 131–143.

Ulberstad, F. (2016). *QbTest technical manual*. Stockholm, Sweden: Qbtech AB.

Vogt, C., & Shameli, A. (2011). Assessments for attention-deficit hyperactivity disorder: Use of objective measurements. *The Psychiatrist Online*, *35*(10), 380–383.

Vogt, C., & Williams, T. (2011). Early identification of stimulant treatment responders, partial responders and non-responders using objective measures in children and adolescents with hyperkinetic disorder. *Child and Adolescent Mental Health*, *16*(3), 144–149.

Zucker, M., Morris, M. K., Ingram, S. M., Morris, R. D., & Bakeman, R. (2002). Concordance of self- and informant ratings of adults' current and childhood attention-deficit/hyperactivity disorder symptoms. *Psychological Assessment*, *14*, 379–389.

How to cite this article: Ulberstad F, Boström H, Chavanon M-L, et al. Objective measurement of attention deficit hyperactivity disorder symptoms outside the clinic using the QbCheck: Reliability and validity. *Int J Methods Psychiatr Res*. 2020;e1822. https://doi.org/10.1002/mpr.1822